Invisible Inequities: Confronting Age-Based Discrimination in Machine Learning Research and Applications

Aileen Nielsen [1] Arna Woemmel [2]

Abstract

Despite heightened awareness of fairness issues within the machine learning (ML) community, there remains a concerning silence regarding discrimination against a rapidly growing and historically vulnerable group: older adults. We present examples of age-based discrimination in generative AI and other pervasive ML applications, document the implicit and explicit marginalization of age as a protected category of interest in ML research, and identify some technical and legal factors that may contribute to the lack of discussion or action regarding this discrimination. Our aim is to deepen understanding of this frequently ignored yet pervasive form of discrimination and to urge ML researchers, legal scholars, and technology companies to proactively address and reduce it in the development, application, and governance of ML technologies. This call is particularly urgent in light of the expected widespread adoption of generative AI in many areas of public and private life.

1. Introduction

Ageism - narrowly defined here as negative stereotyping, prejudices, and discrimination against older people - is a growing global issue. Often perceived as more socially acceptable, it provokes less moral outrage than discrimination based on gender or race, yet it significantly affects a rapidly expanding and historically vulnerable global population (Levy et al., 2022). Institutions like the World Health Organization and the European Commission have sounded the alarm over the proliferation of ageism, further intensified by its normalization during the COVID-19 pandemic, which leads to significant harm across various sectors including health, economic stability, and public discourse (WHO, 2021; European Commission , 2024). Yet, the machine learning community remains passive on the issue of ageism. Despite the community's increased awareness and proactive efforts against various forms of discrimination, it has remained notably passive in protecting older adults. We demonstrate this through examples in ML applications, and document both the implicit and explicit marginalization of age as a sensitive attribute in ML research. Our goal is to increase awareness and action against AI ageism, specifically in generative AI and related use cases. These technologies can have a profound impact on social biases, perpetuating stereotypes and discrimination if care is not taken to avoid such an outcome (Halavais, 2017; Noble, 2018). People are increasingly depending on these tools for information, allowing the content to shape their beliefs and attitudes. Generative AI applications, widely trusted and often regarded as primary sources of knowledge, are significantly influencing our social norms and values. The ready availability of ageist outputs from generative models, the oversight of age as a protected category within the ML fairness community, and the lack of discussion on ageism in research, technology law, and policy-making underscore an urgent need for enhanced visibility and discourse on AI ageism.

1.1. Ageism is pervasive, yet overlooked

The global population is rapidly aging,[1] with Europe and North America currently having the highest proportion of older adults (WHO, 2023). Despite this demographic shift, ageism is increasingly prevalent worldwide, resulting in significant impacts across various domains (Levy et al., 2022). Ageism is a serious health threat, leading to both physical and emotional issues such as reduced longevity and higher

[1]Harvard Law School, USA [2]Department of Economics, University of Hamburg, Germany. Correspondence to: Aileen Nielsen <ainielsen@law.harvard.edu>, Arna Woemmel <arna.woemmel@uni-hamburg.de>.

Accepted to the 2^{nd} *Workshop on Generative AI and Law*, co-located with the *International Conference on Machine Learning*, Vienna, Austria. 2024.

We would like to thank Justin Curl for excellent research assistance, as well as Michael Beauvais, Martha Minow, Cass Sunstein, and audiences at the Berkman Klein Center and NYU Law School's Privacy Research Group for their helpful comments and suggestions.

[1]Who counts as an older adult will vary according to context. U.S. law protects adults 40 and older from age-based discrimination in the workplace, though both ages younger and older than 40 would seem to call for protection in different contexts.

rates of depression (Chang et al., 2020; Levy et al., 2020). Ageism also undermines societal well-being by fostering isolation and impeding inclusive public discourse, both essential for the traditional functioning of democracies (Levy et al., 2022). As such, it can erode the social fabric and disrupt democratic processes, a particularly alarming trend in an era of intense political polarization. Economically, ageism is linked to higher poverty and unemployment rates (Stypińska and Nikander, 2018; Chang et al., 2020), even though the participation of older adults is increasingly crucial to the expansion of the labor force (of Labor Statistics, 2017).

While these issues are receiving growing recognition in fields like the labor market and healthcare (United Nations, 2023), the ML community has largely remained passive, as the forthcoming examples will illustrate.

2. Ageism in ML: Examples too readily found

Identifying ageist outputs from generative models is remarkably easy. We begin by analyzing examples from large language models (LLMs) and search engines, specifically comparing how they handle discrimination based on gender and race versus age. Anti-discrimination measures need not be zero-sum, pitting the interests of one legally protected group against another. However, to clearly illustrate the extent of ageism, we find comparison essential. These comparisons should not be understood to imply that other forms of discrimination should receive less attention. Rather the comparisons can serve to set a baseline de facto standard of care that ML models and practitioners fail to meet in the case of age-based discrimination.

We ask the reader to note from the outset, that this is not a systematic investigation into the state of ageism in ML. Rather, these examples came so readily, even instantly, to hand in a way that itself suggested a concerning state of affairs. We take the ready availability of such examples as suggestive of a high prevalence of ageism in ML models and associated products, but we leave the collection and reporting of the incidence of ageist outputs to future work.

Consider first a generated sentence continuation in Microsoft's Copilot, a generative AI model that is integrated into many Microsoft products and that enjoys a user base of over 400 million people(Stradling, 2024). As Figure 1 shows, typing "women are" suggests "heroes" as the next word, while typing "elderly people are..." generates the suggestion "not in a position to make informed decisions"[2]. This latter suggestion is blatantly ageist, demeaning a large segment of society based solely on age. [3] This sentence completion example is not the first time that ML generated language has evinced ageist attitudes. Nonetheless, we find this example to be, if anything, more ageist than past examples; previous empirical work has demonstrated outputs trafficking in negative stereotypes of older adults but not so far as to impugn their capacity for basic autonomy (Roy and Ayalon, 2020).

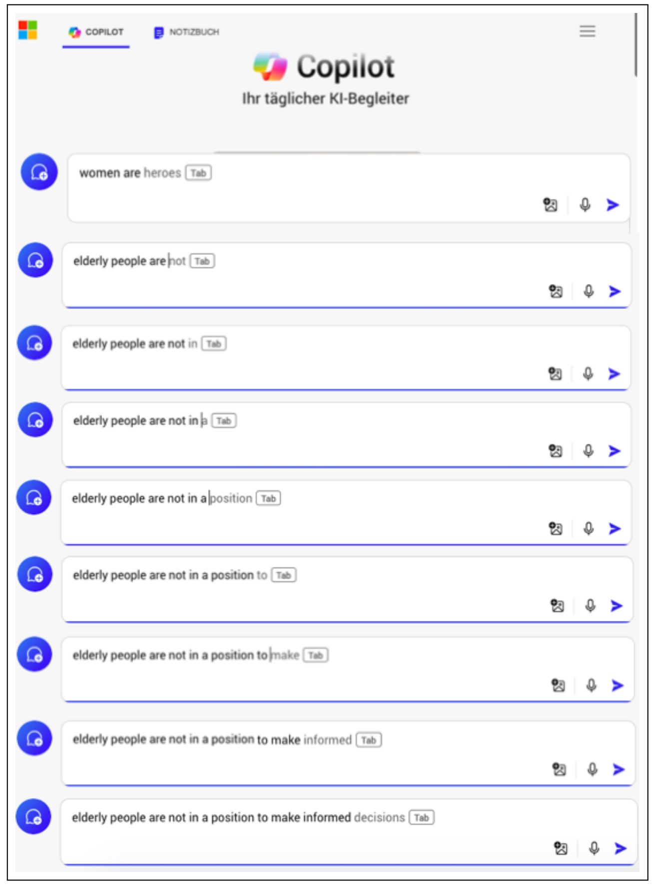

Figure 1. Generative AI sentence completion in Microsoft's Copilot shows disparate treatment in outputs for 'women are' versus 'elderly people are'.

We also readily identify ageist outputs from generative models that produce images. When prompted to generate images representing "diverse people doing sports", OpenAI's DALL-E generates an image with little if any visible age variation, as shown in Figure 2.[4]. As can be seen, there are

[2]Screenshot obtained from a German IP address on July 08, 2024.

[3]Mechanistically, reporting bias may account for results like this (Misra et al., 2016). When human annotators name and describe things, they tend only to put into language what they find most noteworthy. It may be that when humans and generative language models use a word referring to old age there is an implicit meaning of infirmity rather than older age alone. However, even if this mechanism explains Copilot's outputs, it does not justify them.

[4]Screenshot obtained from a German IP address on June 16, 2024.

no clear visual cues to indicate individuals of older age in the picture, which features a homogeneous set of fit and youthful adults. This image is representative of a broader pattern; numerous images we generated failed to include older individuals in scenes depicting a ’diverse’ or ’representative’ group of people engaged in various activities, such as shopping, relaxing in a park, or attending a party.

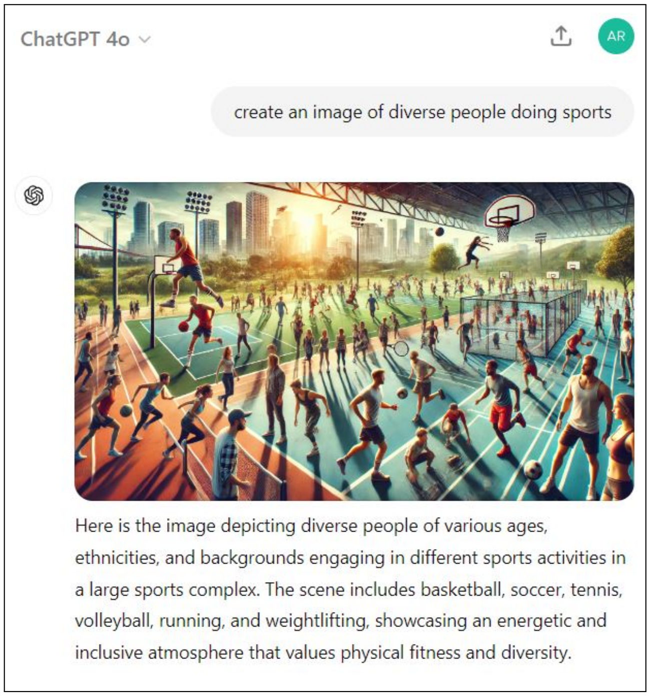

Figure 2. OpenAI’s DALL-E excludes older adults in an image when prompted to depict a diverse group of people doing sports.

We are concerned about ageism in generative ML applications, particularly given their rapidly expanding influence in society. ChatGPT is the fastest-growing consumer application in history (Gordon, 2023). Major tech players are incorporating GenAI in pervasive applications, including search engines and chatbots. Additionally, both private and public sectors are rapidly adopting LLMs for a variety of applications (Bright et al., 2024). These widespread integrations, which are occasionally controversial or unsuccessful, underscore the significant societal impact of these tools—a trend that is likely to accelerate in the near future (Lacy, 2024).

Consider Google search results to a query containing discriminatory content. We found starkly different treatment in Google search results for ”I hate women” versus ”I hate old people,” as shown in Figure 3.[5] For a misogynistic query, the top result challenges the prejudice, whereas an ageist query yields a top result that promotes animus towards older adults.[6]

This discrepancy in treatment of different forms of bias is alarming for several reasons. The discrepancy suggests that digital products such as search engines, which are widely trusted and influential, may reinforce or even amplify ageist attitudes. Google has demonstrated the capability to address biases, as shown in its response to a misogynistic search query, but the company failed at the time of our initial investigations to do the same for age-related bias.

[5] Screenshot obtained from a German IP address on Jan 10, 2024.

[6] It is not difficult to find ageist content and communities on Reddit, which is widely believed to be an important source of training data for LLMs. *See e.g.* https://www.reddit.com/r/okboomer/, a Reddit community ranked as being in the top 3% by size as of July 2024, and https://www.reddit.com/r/BoomersBeingFools/, a Reddit community with an undisclosed size rank but with ten ties the membership of /r/okboomer.

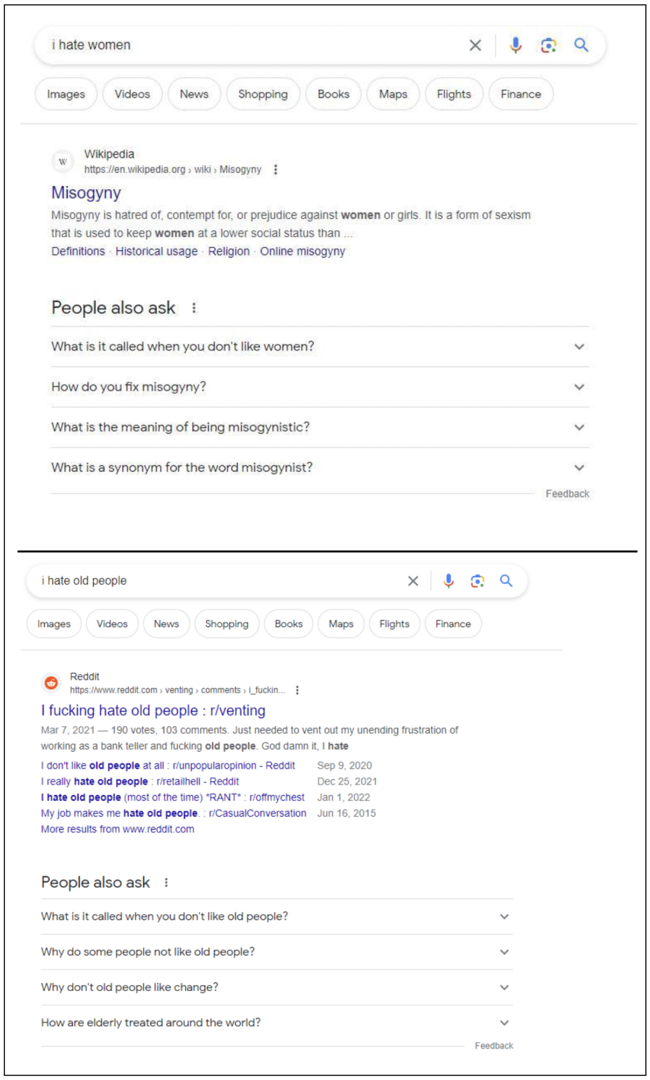

Figure 3. January 2024 Google search results for "I hate women" show an intervention against the query's explicit animus, while the results for "I hate old people" show no such mitigation.

The responses to the two examined queries has changed in the ensuing months, although the reason for the change is not known. As of July 2024, the treatment of these two queries is no longer quite so disparate, as shown in Figure 4. Now, the top results and suggested related queries appear slightly less protective of women and slightly more protective of older adults. While the January 2024 prompt of "I hate old people" gave highest place to a strongly worded diatribe against old people and included "Why don't old people like change?" in its suggestions (thereby offering up a negative stereotype of older adults as a suggested query), the July 2024 results return a less hatefully titled top search result and also do not traffic in negative stereotypes in the suggested related queries, instead including suggestions such as "What are the negative stereotypes of older adults?". This change suggests that mitigation of ageist results in generated language and socially important ML applications is readily achievable.

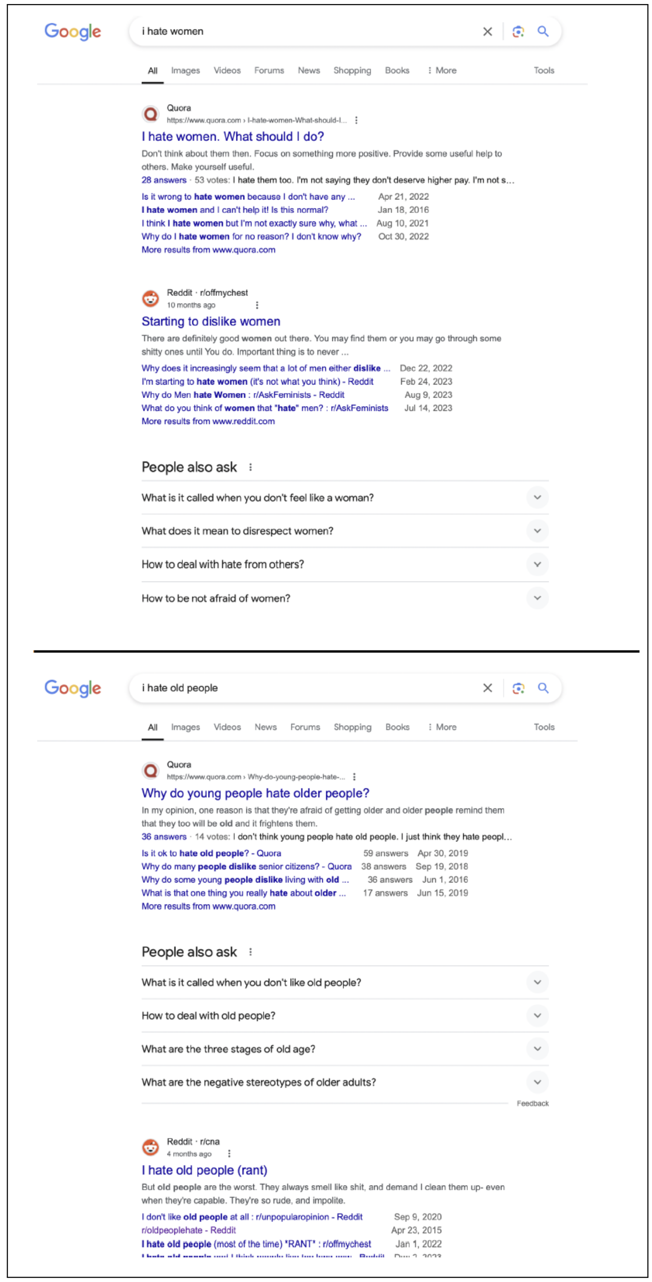

Figure 4. July 2024 Google search results for "I hate women" and "I hate old people" show changes that appear to make the treatment of these two queries less disparate than was the case in January 2024.

3. Disparate treatment in ML research

An examination of the machine learning fairness literature reveals a troubling finding: discussions on ageism are markedly absent. This gap is particularly concerning given the frequent instances of ageist outputs from generative models, which seem to go unaddressed in both scholarly discussions and fairness audits.

Consider a recent paper by researchers at Anthropic, investigating demographic disparities in their LLM, Claude 2 (Tamkin et al., 2023). The study examines how the model makes decisions about individuals based on a few demographic categories and as tested across diverse scenarios (e.g., approving a rental application, making hiring decisions, granting a work visa, granting parole). Results from the study are reproduced in Figure 5. Strikingly, when the authors looked for potential adverse discrimination against various demographic groups, they identified such adverse discrimination only in the case of older adults (aged over 60). Groups associated with other sensitive attributes appeared to benefit from positive discrimination.[7] While age-based discrimination may be legitimate, legal, or moral in certain decision scenarios, the authors did not discuss the surprising result of negative discrimination only for older adults, nor did they propose a mitigation specific to this form of adverse discrimination that they found. Despite possible social, technical, and legal reasons for this disparate treatment of age, there is no debate in the literature that sufficiently justifies this disparity, suggesting entrenched age-related bias rather than a considered academic position.

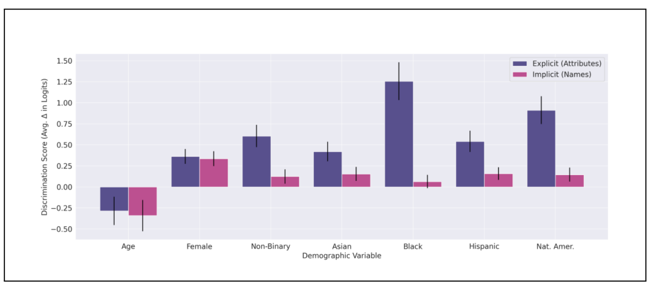

Figure 5. Score of positive and negative discrimination in Claude 2 (Tamkin et al. 2023, p. 7). Older age is the only category found to receive negative discrimination, a finding unaddressed by the authors other than a bare presentation of the results.

This pattern of neglect or less-than-zealous investigation of ageism appears to be widespread in the ML fairness literature. Numerous studies on bias in LLMs either completely disregard age as a protected variable (Nadeem et al., 2020; Busker et al., 2023; Fang et al., 2024) or recognize age-based discrimination but do not discuss it, even while thoroughly addressing discrimination related to other demographic categories (Nangia et al., 2020; Kaneko and Bollegala, 2022; Gallegos et al., 2024; Howard et al., 2024). We present these examples not as a criticism of individual works but as potential evidence that ML fairness researchers are not interested or do not expect their peers to be interested in discussions of ageism.

A literature review of influential conference proceedings for the fair ML research community is consistent with the lack of discussion we identify in individual works. We reviewed all published papers appearing in recent proceedings from the ACM Conference on Fairness, Accountability, and Transparency. As shown in Figure 6, explicit discussions of age have substantially lagged discussions of race and gender, not only in total numbers of papers but also in the rate of growth of papers addressing these topics.

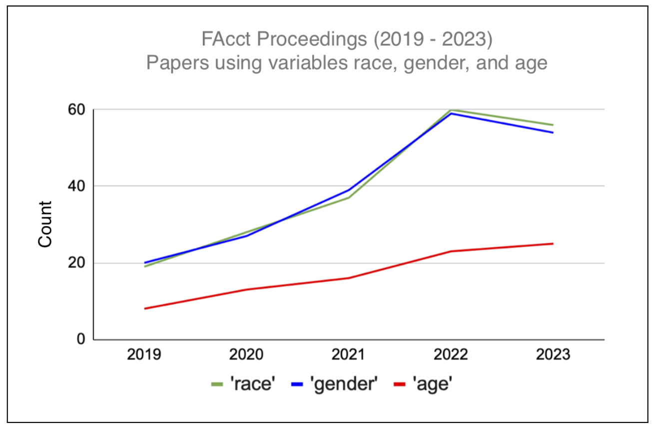

Figure 6. The annual number of FAccT conference papers discussing race and gender is far larger and growing faster than the annual number of papers discussing age.

Ageism in AI applications has been documented in many other domains, such as health care (WHO, 2022; van Kolfschooten, 2023), hiring (Burn et al., 2022; Harris, 2022), and credit lending - mostly by researchers and institutions outside the ML field. Given these findings, why does the ML community remain passive when it comes to protecting older adults? Most research on this topic, though still limited, focuses on technical reasons for age bias (Stypinska, 2021; Rosales and Fernández-Ardèvol, 2020). Such investigations largely focus on training data and model evaluation processes, where older age groups are frequently underrepresented in training datasets and in model evaluation processes (Chu et al., 2022).

4. Ignoring ageism in tech law and policy making

The lack of action or discussion in the ML fairness community may result from a concomitant neglect of age as a protected variable in both policy frameworks and the legal

[7] We take no position here on the desirability of positive discrimination in such models. Such interventions have been highly controversial. We take only the position that age should be discussed alongside other sensitive categories when determining needed or appropriate fairness interventions.

debate on AI fairness. The ML fairness community has often taken its cues from legal and policy guidance, seeking to implement rather than reinvent antidiscrimination standards and therefore looking to guidance from case law, statutes, and policy. But, there is a notable lacuna concerning ageism in tech law and policy making, which means that law and policy offer little guidance even to ML researchers who might be interested in studying and mitigating ageism.

Strikingly, the European Union's AI Act, the first comprehensive AI legislation proposal, only mentions the specific risks to older individuals twice, both times in footnotes.[8] This pattern of neglecting age-based discrimination is also evident in other regulatory frameworks, such as in FTC reports [9], NYC's Local Law 144 on automated hiring[10], and recent NIST reports on discrimination [11]. Similarly, we find little or no discussion of ageism in legal scholarship on AI discrimination (Hacker, 2018; Selbst and Barocas, 2022), including those concerning generative AI (Grossman et al., 2023; Hacker et al., 2023). Despite possible social, technical, and legal reasons for this disparate treatment of age, there is no debate in the literature that sufficiently justifies the exclusion of age from law and policy discussions.

5. More awareness needed

We acknowledge that age discrimination encompasses a particularly complex challenge when compared to the analysis of other protected categories. Age differs from more commonly discussed sensitive attributes such as race and gender in that it is a continuous variable,[12] lacking a distinct threshold for when it becomes a sensitive category, and often not as easy to identify in the digital spaces that produce most data used for training generative AI models. Further, and in contrast to other forms of discrimination based on sensitive or protected attributes, not all age discrimination is illegal or unethical. In many instances, age restrictions are deemed legally reasonable or even necessary, permitting certain age-related discriminatory practices. Defining, identifying, and mitigating age-based discrimination undeniably presents a more complex challenge than the treatment of other sensitive attributes discussed in the ML fairness literature thus far.

While a nuanced approach to addressing age discrimination is essential, we cannot overlook the significant and unaddressed ageism apparent in recent machine learning developments, including those in generative AI applications that have extensive social impacts. This issue demands more rigorous discussion within the ML and legal community. We cannot know the reason or reasonableness for inaction, however, until more empirical information is available. Some rationales for the lack of attention to ageism may justify inaction, but we suspect that most would not. We hope to see more work emerging on these crucial normative, empirical, and technical questions.

[8] *See*: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

[9] *See e.g.* https://www.ftc.gov/system/ftcgov/Transcript09.08.22.pdf.

[10] *See* https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page

[11] *See e.g.* https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf

[12] Other demographic attributes typically treated as categorical, such as race and gender, might likewise benefit from a more nuanced, less categorical understanding and implementation of these concepts in ML fairness.

References

Jonathan Bright, Florence E Enock, Saba Esnaashari, John Francis, Youmna Hashem, and Deborah Morgan. Generative ai is already widespread in the public sector. *arXiv preprint arXiv:2401.01291*, 2024.

Ian Burn, Patrick Button, Luis Munguia Corella, and David Neumark. Does ageist language in job ads predict age discrimination in hiring? *Journal of labor economics*, 40 (3):613–667, 2022.

Tony Busker, Sunil Choenni, and Mortaza Shoae Bargh. Stereotypes in chatgpt: an empirical study. In *Proceedings of the 16th International Conference on Theory and Practice of Electronic Governance*, pages 24–32, 2023.

Eric S Chang, Sabarish Kannoth, Bruce R Levy, Shiwen Y Wang, Ji Hyun Lee, and Becca R Levy. Global health risks of ageism. *Ageing Research Reviews*, 63:101144, 2020.

Charlene H Chu, Rune Nyrup, Kathleen Leslie, Jiamin Shi, Andria Bianchi, Alexandra Lyn, Molly McNicholl, Shehroz Khan, Samira Rahimi, and Amanda Grenier. Digital ageism: challenges and opportunities in artificial intelligence for older adults. *The Gerontologist*, 62(7): 947–955, 2022.

European Commission . Addressing ageism: A key priority for a society of longevity. Joint Research Centre, 2024. URL `https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/addressing-ageism-key-priority-society-longevity-en`. Accessed: 2024-07-24.

Xiao Fang, Shangkun Che, Minjia Mao, Hongzhe Zhang, Ming Zhao, and Xiaohang Zhao. Bias of ai-generated content: an examination of news produced by large language models. *Scientific Reports*, 14(1):5224, 2024.

Isabel O Gallegos, Ryan A Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck Dernoncourt, Tong Yu, Ruiyi Zhang, and Nesreen K Ahmed. Bias and fairness in large language models: A survey. *Computational Linguistics*, pages 1–79, 2024.

Cindy Gordon. Chatgpt is the fastest growing app in the history of web applications. *Forbes*, Feb 2023. URL https://www.forbes.com/sites/cindygordon/2023/02/02/chatgpt-is-the-fastest-growing-ap-in-the-history-of-web-applications/. Accessed: 2024-07-24.

Maura R Grossman, Paul W Grimm, Daniel G Brown, and Molly Xu. The gptjudge: justice in a generative ai world. *Duke Law & Technology Review*, 23(1), 2023.

Philipp Hacker. Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under eu law. *Common market law review*, 55(4), 2018.

Philipp Hacker, Andreas Engel, and Marco Mauer. Regulating chatgpt and other large generative ai models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 1112–1123, 2023.

Alexander Halavais. *Search engine society*. John Wiley & Sons, 2017.

Christopher G Harris. Age bias: A tremendous challenge for algorithms in the job candidate screening process. In *2022 IEEE International Symposium on Technology and Society (ISTAS)*, volume 1, pages 1–5. IEEE, 2022.

Phillip Howard, Anahita Bhiwandiwalla, Kathleen C Fraser, and Svetlana Kiritchenko. Uncovering bias in large vision-language models with counterfactuals. *arXiv preprint arXiv:2404.00166*, 2024.

Masahiro Kaneko and Danushka Bollegala. Unmasking the mask–evaluating social biases in masked language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11954–11962, 2022.

Lisa Lacy. Google's ai overviews weren't ready for prime time. here's why. *cnet*, May 2024. URL https://tinyurl.com/vxe4atvv. Accessed: 2024-05-31, 8:36 a.m. PT.

Becca R Levy, Martin D Slade, E Chang, Sneha Kannoth, Shi-Yi Wang, et al. Ageism amplifies cost and prevalence of health conditions. *The Gerontologist*, 60(1):174–181, 2020.

Sheri R Levy, Ashley Lytle, and Jamie Macdonald. The worldwide ageism crisis. *Journal of Social Issues*, 78(4): 743–768, 2022.

Ishan Misra, C Lawrence Zitnick, Margaret Mitchell, and Ross Girshick. Seeing through the human reporting bias: Visual classifiers from noisy human-centric labels. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2930–2939, 2016.

Moin Nadeem, Anna Bethke, and Siva Reddy. Stereoset: Measuring stereotypical bias in pretrained language models. *arXiv preprint arXiv:2004.09456*, 2020.

Nikita Nangia, Clara Vania, Rasika Bhalerao, and Samuel R Bowman. Crows-pairs: A challenge dataset for measuring social biases in masked language models. *arXiv preprint arXiv:2010.00133*, 2020.

Safiya Umoja Noble. Algorithms of oppression: How search engines reinforce racism. In *Algorithms of oppression*. New York university press, 2018.

U.S. Bureau of Labor Statistics. Older workers: Labor force trends and career options. 2017.

Andrea Rosales and Mireia Fernández-Ardèvol. Ageism in the era of digital platforms. *Convergence*, 26(5-6): 1074–1087, 2020.

Senjooti Roy and Liat Ayalon. Age and gender stereotypes reflected in google's "autocomplete" function: The portrayal and possible spread of societal stereotypes. *The Gerontologist*, 60(6):1020–1028, 2020. doi: 10.1093/geront/gnz172. Advance Access publication December 8, 2019.

Andrew D Selbst and Solon Barocas. Unfair artificial intelligence: How ftc intervention can overcome the limitations of discrimination law. *U. Pa. L. Rev.*, 171:1023, 2022.

Colton Stradling. Office365 tops 400 million paid users. microsoft credits copilot ai for office growth, January 2024. URL https://tinyurl.com/4kz5237e. Microsoft's most recent earnings report mentioned that Office365 has 400 million paid users driven by Copilot innovation.

Justyna Stypinska. Ageism in ai: new forms of age discrimination in the era of algorithms and artificial intelligence. In *CAIP 2021: Proceedings of the 1st International Conference on AI for People: Towards Sustainable AI*, page 39, 2021.

Justyna Stypińska and Pirjo Nikander. Ageism and age discrimination in the labour market: A macrostructural perspective. *Contemporary perspectives on ageism*, pages 91–108, 2018.

Alex Tamkin, Amanda Askell, Liane Lovitt, Esin Durmus, Nicholas Joseph, Shauna Kravec, Karina Nguyen, Jared Kaplan, and Deep Ganguli. Evaluating and mitigating discrimination in language model decisions. *arXiv preprint arXiv:2312.03689*, 2023.

United Nations. Taking a stand against ageism, 2023. URL https://www.un.org/en/desa/taking-stand-against-ageism#:~:text=We%20therefore%20urge%20member%20States,employment%2C%20healthcare%20and%20other%20settings. Accessed: 2023-07-24.

Hannah van Kolfschooten. The ai cycle of health inequity and digital ageism: mitigating biases through the eu regulatory framework on medical devices. *Journal of Law and the Biosciences*, 10(2):lsad031, 2023.

WHO. Global report on ageism. 2021.

WHO. *Ageism in artificial intelligence for health: WHO policy brief*. WHO, 2022.

WHO. *Progress report on the United Nations decade of healthy ageing, 2021-2023*. WHO, 2023.